New Methods for Modeling Large-Scale Biochemical Networks

Will Heuett
University of Washington
Department of Applied Mathematics

15 December 2005

Outline

I will give an overview of the current state of systems biology and discuss the contributions I have made in this field.

- Background and motivation for the new modeling methods.
- The grand canonical model.
- Stochastic simulation algorithms.
- Stoichiometric constraints-based optimization approaches.
- Conclusions.

Outline

- Background and motivation for the new modeling methods.
- The grand canonical model.
- Stochastic simulation algorithms.
- Stoichiometric constraints-based optimization approaches.
- Conclusions.

Systems Biology

Systems biology is a method of study where the goal is to understand how the parts of a biological system interact to yield the behavior of the system as a coherent whole (Iyengar et al., 2004).

Systems Biology

Systems biology is a method of study where the goal is to understand how the parts of a biological system interact to yield the behavior of the system as a coherent whole (Iyengar et al., 2004).

- Different levels of study: cells, organs, organisms, and populations.

Systems Biology

Systems biology is a method of study where the goal is to understand how the parts of a biological system interact to yield the behavior of the system as a coherent whole (Iyengar et al., 2004).

- Different levels of study: cells, organs, organisms, and populations.
- Ultimate goal is to develop a complete model of a living organism.

Systems Biology

Systems biology is a method of study where the goal is to understand how the parts of a biological system interact to yield the behavior of the system as a coherent whole (Iyengar et al., 2004).

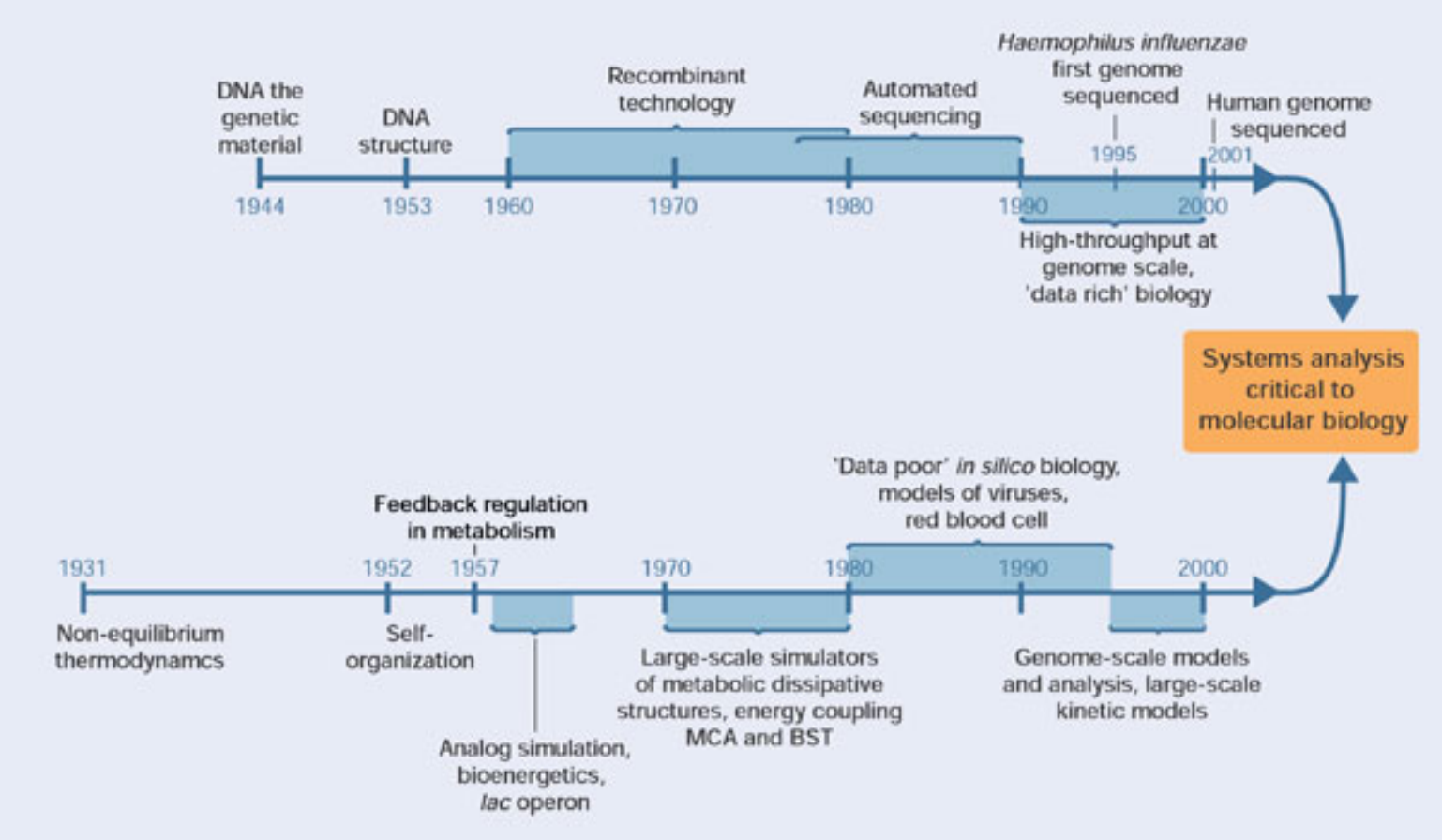
- Different levels of study: cells, organs, organisms, and populations.
- Ultimate goal is to develop a complete model of a living organism.
- The field is comprised of researchers from several different backgrounds.

Systems Biology

Systems biology is a method of study where the goal is to understand how the parts of a biological system interact to yield the behavior of the system as a coherent whole (Iyengar et al., 2004).

- Different levels of study: cells, organs, organisms, and populations.
- Ultimate goal is to develop a complete model of a living organism.
- The field is comprised of researchers from several different backgrounds.
- So mathematical models need to be accessible to experimentalists.

Scaling-Up to Systems Biology



(Westerhoff and Palsson, 2004)

Aspects of Biochemical Networks

- Biochemical networks can be very large and complex $\Longrightarrow$
  many models are developed for small subsets of the overall network.

Aspects of Biochemical Networks

- Biochemical networks can be very large and complex $\Longrightarrow$
  many models are developed for small subsets of the overall network.

- Living organisms interact with their environment $\Longrightarrow$
  difference between closed and open (living and dead) systems.

Aspects of Biochemical Networks

- Biochemical networks can be very large and complex $\Longrightarrow$
  many models are developed for small subsets of the overall network.

- Living organisms interact with their environment $\Longrightarrow$
  difference between closed and open (living and dead) systems.

- There are many other factors to consider when building a model $\Longrightarrow$
  mutations or gene deletions, dynamic responses to perturbations, effects of small numbers of molecules.

Aspects of Biochemical Networks

- Biochemical networks can be very large and complex $\Longrightarrow$
  many models are developed for small subsets of the overall network.
- Living organisms interact with their environment $\Longrightarrow$
  difference between closed and open (living and dead) systems.
- There are many other factors to consider when building a model $\Longrightarrow$
  mutations or gene deletions, dynamic responses to perturbations, effects of small numbers of molecules.
- We would like to be able to do *in silico* experiments with our models.

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

- The Law of Mass Action:
  - Wilhelmy (1850) measured the velocity of mutarotation of simple sugars.
  - Waage and Guldberg (1864, 1867) assumed reversibility of each elementary reaction and identified the 'forward' and 'reverse' rates.
  - Harcourt and Esson (1866) discovered the law independently.

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

- The Law of Mass Action:
  - Wilhelmy (1850) measured the velocity of mutarotation of simple sugars.
  - Waage and Guldberg (1864, 1867) assumed reversibility of each elementary reaction and identified the 'forward' and 'reverse' rates.
  - Harcourt and Esson (1866) discovered the law independently.
- Michaelis–Menten Enzyme Kinetics:
  - Named after Michaelis and Menten (1913).

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

- The Law of Mass Action:
  - Wilhelmy (1850) measured the velocity of mutarotation of simple sugars.
  - Waage and Guldberg (1864, 1867) assumed reversibility of each elementary reaction and identified the 'forward' and 'reverse' rates.
  - Harcourt and Esson (1866) discovered the law independently.
- Michaelis–Menten Enzyme Kinetics:
  - Named after Michaelis and Menten (1913).
- Nonequilibrium Thermodynamics:
  - Influential work by Onsager (1931) and Hill (1989).

The Law of Mass Action

For a system involving M reactions and N chemical species with j^{th} reaction

$$\nu_1^j X_1 + \nu_2^j X_2 + \cdots + \nu_N^j X_N \underset{k_-^j}{\overset{k_+^j}{\rightleftharpoons}} \kappa_1^j X_1 + \kappa_2^j X_2 + \cdots + \kappa_N^j X_N,$$

the law of mass action gives

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{M} (\kappa_i^j - \nu_i^j)(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}).$$

The Law of Mass Action

For a system involving M reactions and N chemical species with j^{th} reaction

$$\nu_1^j X_1 + \nu_2^j X_2 + \cdots + \nu_N^j X_N \underset{k_-^j}{\overset{k_+^j}{\rightleftharpoons}} \kappa_1^j X_1 + \kappa_2^j X_2 + \cdots + \kappa_N^j X_N,$$

the law of mass action gives

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{M} (\kappa_i^j - \nu_i^j)(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}).$$

A closed system will go to equilibrium, whereas an open system will go to a nonequilibrium steady state (NESS).

Detailed Balance

When in equilibrium, the forward and reverse fluxes are equal

$$k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} = k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j},$$

which yields

$$\frac{k_+^j}{k_-^j} = \frac{x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}}{x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j}} = K_{eq}^j,$$

and, for a closed loop of reactions $j_1 \rightarrow j_2 \rightarrow \cdots \rightarrow j_z \rightarrow j_1$,

$$\frac{k_+^{j_1} k_+^{j_2} \cdots k_+^{j_z}}{k_-^{j_1} k_-^{j_2} \cdots k_-^{j_z}} = 1.$$

Open, Living Systems

Starting with the original mass-action kinetics

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{M} (\kappa_i^j - \nu_i^j)(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}),$$

the detailed balance conditions can be broken by incorporating external input and output fluxes

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{M} (\kappa_i^j - \nu_i^j)(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}) + J_i^{ext}$$

or concentration clamping

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{M} (\kappa_i^j - \nu_i^j)(k_+^j c_0^{\nu_0^j} c_{N+1}^{\nu_{N+1}^j} x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j c_0^{\kappa_0^j} c_{N+1}^{\kappa_{N+1}^j} x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}).$$

Michaelis–Menten Enzyme Kinetics

For enzyme-catalyzed reactions, represented as

$$S + E \underset{k_-^1}{\overset{k_+^1}{\rightleftharpoons}} SE \xrightarrow{k_+^2} E + P,$$

we can make the quasi-steady-state assumption that

$$\frac{d}{dt}(se) \approx 0 \implies (se) = \frac{k_+^1 s \cdot e}{k_-^1 + k_+^2} = \frac{s \cdot e}{K_M},$$

where $K_{M,s}$ is known as the Michaelis–Menten rate constant.

Michaelis–Menten Enzyme Kinetics

For enzyme-catalyzed reactions, represented as

$$S + E \underset{k_-^1}{\overset{k_+^1}{\rightleftharpoons}} SE \xrightarrow{k_+^2} E + P,$$

we can make the quasi-steady-state assumption that

$$\frac{d}{dt}(se) \approx 0 \implies (se) = \frac{k_+^1 s \cdot e}{k_-^1 + k_+^2} = \frac{s \cdot e}{K_M},$$

where $K_{M,s}$ is known as the Michaelis–Menten rate constant.

This yields

$$\frac{ds}{dt} = -\frac{dp}{dt} = -\frac{k_+^2 e_0 \cdot s}{K_{M,s} + s} = -\frac{V_{max}^+ \, s}{K_{M,s} + s}.$$

Reversible Michaelis–Menten Enzyme Kinetics

For

$$S + E \underset{k_-^1}{\overset{k_+^1}{\rightleftharpoons}} SE \underset{k_-^2}{\overset{k_+^2}{\rightleftharpoons}} E + P,$$

we get

$$\frac{ds}{dt} = -\frac{dp}{dt} = -\frac{V_{max}^+ \frac{s}{K_{M,s}} - V_{max}^- \frac{p}{K_{M,p}}}{1 + \frac{s}{K_{M,s}} + \frac{p}{K_{M,p}}},$$

where

$$K_{M,s} = \frac{k_-^1 + k_+^2}{k_+^1} \quad \text{and} \quad K_{M,p} = \frac{k_-^1 + k_+^2}{k_-^2}.$$

Nonequilibrium Thermodynamics

The chemical potential of a species is given by

$$\mu_i = \mu_i^o + RT \ln x_i,$$

from which we get the reaction potential, given by

$$\Delta\mu^j = RT \ln \left(\frac{k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}}{k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j}} \right).$$

Nonequilibrium Thermodynamics

The chemical potential of a species is given by

$$\mu_i = \mu_i^o + RT \ln x_i,$$

from which we get the reaction potential, given by

$$\Delta\mu^j = RT \ln \left(\frac{k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}}{k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j}} \right).$$

It is easy to show that

$$-\left(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j} \right) \Delta\mu^j \geq 0$$

and, for a closed loop of reactions $j_1 \to j_2 \to \cdots \to j_z \to j_1$,

$$\Delta\mu^{j_1} + \Delta\mu^{j_2} + \cdots + \Delta\mu^{j_z} = 0.$$

Limitations of Classical Methods

- Typically, the biochemical networks of interest are very large and complex. As a result, it is extremely difficult to solve for analytic solutions of the models. If they can be found, general solutions for particular types of networks can be very important.

Limitations of Classical Methods

- Typically, the biochemical networks of interest are very large and complex. As a result, it is extremely difficult to solve for analytic solutions of the models. If they can be found, general solutions for particular types of networks can be very important.

- It is quite common that the chemical species of a system are only available in small numbers. In such cases, deterministic models are not valid and stochastic models are needed to capture the effects of intrinsic fluctuations.

Limitations of Classical Methods

- Typically, the biochemical networks of interest are very large and complex. As a result, it is extremely difficult to solve for analytic solutions of the models. If they can be found, general solutions for particular types of networks can be very important.

- It is quite common that the chemical species of a system are only available in small numbers. In such cases, deterministic models are not valid and stochastic models are needed to capture the effects of intrinsic fluctuations.

- Experimentalists are limited in the amount of information they can gather and, in most cases, it is not possible to obtain detailed kinetic-rate information. Therefore, methods that avoid having to know this information are needed.

Outline

- Background and motivation for the new modeling methods.
- The grand canonical model.
- Stochastic simulation algorithms.
- Stoichiometric constraints-based optimization approaches.
- Conclusions.

Stochastic Chemical Kinetics

- The mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.

Stochastic Chemical Kinetics

- The mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.

- Kurtz (1972) proved, by properly taking the volume of the system into account, that the deterministic ODE model is the infinite limit of the Markov chain model.

Stochastic Chemical Kinetics

- The mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.
- Kurtz (1972) proved, by properly taking the volume of the system into account, that the deterministic ODE model is the infinite limit of the Markov chain model.
- There are *state-tracking* approaches, such as the single-molecule NESS studies from the work of Hill (1989), and there are *number-tracking* approaches, such as the simulation methods developed by Gillespie (1977).

Stochastic Chemical Kinetics

- The mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.
- Kurtz (1972) proved, by properly taking the volume of the system into account, that the deterministic ODE model is the infinite limit of the Markov chain model.
- There are *state-tracking* approaches, such as the single-molecule NESS studies from the work of Hill (1989), and there are *number-tracking* approaches, such as the simulation methods developed by Gillespie (1977).
- Before now, analytic solutions were not available for models that considered explicit material exchange between the system and its surroundings. Using the grand canonical model, we can obtain such solutions for unimolecular reaction networks.

The Closed Network Model

Consider an irreducible system of monomolecular biochemical reactions involving N reactants. The dynamics of a single molecule can be modeled as a Markov jump process or random walk, where

$$\frac{dp_i(t)}{dt} = \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(p_j(t) q_{j,i} - p_i(t) q_{i,j} \right) .$$

If there are a total of n molecules in the closed system, then we have the joint probability

$$P(n_1, n_2, \ldots, n_N, t) = \frac{n!}{n_1! n_2! \cdots n_N!} \left(p_1(t) \right)^{n_1} \left(p_2(t) \right)^{n_2} \cdots \left(p_N(t) \right)^{n_N} ,$$

i.e., the multinomial distribution.

Closed Network $\Longleftrightarrow$ Equilibrium

If the system is closed to its surroundings, it will achieve equilibrium where each reaction must be detailed balanced. Let π_i represent the probability of being in state i at equilibrium, then

$$\frac{\pi_i}{\pi_j} = \frac{q_{j,i}}{q_{i,j}} = K_{eq}$$

and, for a closed loop of reactions $i_0 \rightarrow i_1 \rightarrow \cdots \rightarrow i_m \rightarrow i_0$,

$$\frac{q_{i_0,i_1} q_{i_1,i_2} \cdots q_{i_m,i_0}}{q_{i_1,i_0} q_{i_2,i_1} \cdots q_{i_0,i_m}} = 1.$$

The detailed balance conditions can be broken for open systems by fixing external concentrations that are typically absorbed into pseudo-first-order transition rate constants.

The Grand Canonical Model

Suppose there is an observer doing work on the system by keeping the number of molecules in state 0 equal to n_0 and the number in state $N+1$ equal to n_{N+1}. The total number of molecules in this system will fluctuate with the expected number of molecules, $\langle n_i(t) \rangle$, at time t in state i satisfying

$$\frac{d \langle n_i(t) \rangle}{dt} = \sum_{\substack{j=0 \\ j \neq i}}^{N+1} \left(\langle n_j(t) \rangle \, q_{j,i} - \langle n_i(t) \rangle \, q_{i,j} \right)$$

$\forall i \in \{1, 2, \ldots, N\}$ with $\langle n_0(t) \rangle = n_0$ and $\langle n_{N+1}(t) \rangle = n_{N+1}$. The system is assumed to be empty initially, so the initial conditions are

$$\text{ICs:} \quad \langle n_i(0) \rangle = 0, \quad \forall i \in \{1, 2, \ldots, N\}.$$

This system is actively exchanging material with its surroundings and will go to a NESS.

Variances and Covariances of the Open System

Let $\langle n_i \rangle^*$ be the number of molecules in state i when the system is in NESS and $\Delta n_i \triangleq n_i - \langle n_i \rangle^*$. It can be shown that the variances and covariances must satisfy

$$\frac{d\left\langle (\Delta n_i(t))^2 \right\rangle}{dt} = \sum_{\substack{j=0 \\ j \neq i}}^{N+1} \left[\left(\langle n_j(t) \rangle + 2 \left\langle \Delta n_i(t) \Delta n_j(t) \right\rangle \right) q_{j,i} \right.$$

$$\left. + \left(\langle n_i(t) \rangle - 2 \left\langle (\Delta n_i(t))^2 \right\rangle \right) q_{i,j} \right]$$

$$\frac{d\left\langle \Delta n_i(t) \Delta n_j(t) \right\rangle}{dt} = \sum_{\substack{k=0 \\ k \neq i}}^{N+1} \left(\left\langle \Delta n_j(t) \Delta n_k(t) \right\rangle q_{k,i} - \left\langle \Delta n_i(t) \Delta n_j(t) \right\rangle q_{i,k} \right)$$

$$+ \sum_{\substack{k=0 \\ k \neq j}}^{N+1} \left(\left\langle \Delta n_i(t) \Delta n_k(t) \right\rangle q_{k,j} - \left\langle \Delta n_i(t) \Delta n_j(t) \right\rangle q_{j,k} \right)$$

$$- \langle n_i(t) \rangle q_{i,j} - \langle n_j(t) \rangle q_{j,i}.$$

The Analytic Solution of the Open System

Using

$$\frac{d\left\langle n_i(t)\right\rangle}{dt} = \sum_{\substack{j=0 \\ j\neq i}}^{N+1} \left(\left\langle n_j(t)\right\rangle q_{j,i} - \left\langle n_i(t)\right\rangle q_{i,j}\right),$$

we obtain the solution of the chemical master equation for the open system, which is given by the joint probability

$$P(n_1, n_2, \ldots, n_N, t) = \prod_{i=1}^{N} \left[\frac{\left\langle n_i(t)\right\rangle^{n_i}}{n_i!} e^{-\left\langle n_i(t)\right\rangle}\right].$$

This solution shows that the numbers of molecules in each state are independent and the corresponding random variables each have Poisson distributions.

The Homogeneous System

The nonhomogeneous system of equations can be rewritten as a homogeneous system of equations by substituting $n_i(t) = \Delta n_i(t) + \langle n_i \rangle^*$. In matrix form, the resulting system is written as

$$\frac{d\langle \Delta \mathbf{n}(t) \rangle}{dt} = \mathrm{Q}\langle \Delta \mathbf{n}(t) \rangle ,$$

where

$$\mathrm{Q} = \begin{pmatrix} -\sum_{j=0}^{N+1} q_{1,j} & q_{2,1} & q_{3,1} & \cdots & q_{N,1} \\ q_{1,2} & -\sum_{j=0}^{N+1} q_{2,j} & q_{3,2} & \cdots & q_{N,2} \\ \vdots & & \ddots & & \vdots \\ q_{1,N-1} & q_{2,N-1} & \cdots & -\sum_{j=0}^{N+1} q_{N-1,j} & q_{N,N-1} \\ q_{1,N} & q_{2,N} & \cdots & q_{N-1,N} & -\sum_{j=0}^{N+1} q_{N,j} \end{pmatrix} .$$

Correlation Functions and Reaction Fluxes

Using the solution of the homogeneous system, we get the autocorrelation and cross-correlation functions

$$\langle \Delta n_i(t) \Delta n_i(0) \rangle = e^{\mathsf{Q}t}_{i,i} \langle n_i \rangle^* = \sum_{k=1}^{N} \left(\mathsf{V}_{i,k} e^{\lambda_k t} \mathsf{V}^{-1}_{k,i} \langle n_i \rangle^* \right)$$

$$\langle \Delta n_j(t) \Delta n_i(0) \rangle = e^{\mathsf{Q}t}_{j,i} \langle n_i \rangle^* = \sum_{k=1}^{N} \left(\mathsf{V}_{j,k} e^{\lambda_k t} \mathsf{V}^{-1}_{k,i} \langle n_i \rangle^* \right).$$

From the cross-correlation functions, we can get the NESS flux of a reaction, as well as the one-way fluxes.

$$\lim_{t \to 0} \frac{\langle \Delta n_j(t) \Delta n_i(0) \rangle - \langle \Delta n_i(t) \Delta n_j(0) \rangle}{t} = q_{i,j} \langle n_i \rangle^* - q_{j,i} \langle n_j \rangle^*.$$

A Simple Example

Consider

$$0 \underset{q_{1,0}}{\overset{q_{0,1}}{\rightleftharpoons}} 1 \underset{q_{2,1}}{\overset{q_{1,2}}{\rightleftharpoons}} 2 \xrightarrow{q_{2,3}} 3,$$

for which we have the master equation

$$\begin{aligned}\frac{dP(n_1,n_2,t)}{dt} &= P(n_1-1,n_2,t)q_{0,1}n_0 + P(n_1,n_2+1,t)q_{2,3}(n_2+1)\\ &+ P(n_1+1,n_2,t)q_{1,0}(n_1+1)\\ &+ P(n_1+1,n_2-1,t)q_{1,2}(n_1+1)\\ &+ P(n_1-1,n_2+1,t)q_{2,1}(n_2+1)\\ &- P(n_1,n_2,t)\left(q_{0,1}n_0 + (q_{1,0}+q_{1,2})n_1 + (q_{2,1}+q_{2,3})n_2\right),\end{aligned}$$

with solution

$$P(n_1,n_2,t) = \frac{\langle n_1(t)\rangle^{n_1}}{n_1!}e^{-\langle n_1(t)\rangle}\frac{\langle n_2(t)\rangle^{n_2}}{n_2!}e^{-\langle n_2(t)\rangle}.$$

A Simple Example

At NESS,

$$\langle n_1 \rangle^* = \frac{q_{0,1}(q_{2,1}+q_{2,3})n_0}{q_{1,0}(q_{2,1}+q_{2,3})+q_{1,2}q_{2,3}}$$
$$\langle n_2 \rangle^* = \frac{q_{0,1}q_{1,2}n_0}{q_{1,0}(q_{2,1}+q_{2,3})+q_{1,2}q_{2,3}},$$

$$\left\langle (\Delta n_1)^2 \right\rangle^* = \langle n_1 \rangle^*, \quad \left\langle (\Delta n_2)^2 \right\rangle^* = \langle n_2 \rangle^*, \quad \langle \Delta n_1 \Delta n_2 \rangle^* = 0,$$

and

$$P(n_1,n_2) = \frac{\left(\langle n_1 \rangle^*\right)^{n_1}}{n_1!} e^{-\langle n_1 \rangle^*} \frac{\left(\langle n_2 \rangle^*\right)^{n_2}}{n_2!} e^{-\langle n_2 \rangle^*}.$$

A Simple Example

The autocorrelation and cross-correlation function are

$$\langle \Delta n_1(0)\Delta n_1(t)\rangle = \frac{\langle n_1\rangle^*}{\lambda_1-\lambda_2}\left((\lambda_1+q_{1,0}+q_{1,2})e^{\lambda_2 t}-(\lambda_2+q_{1,0}+q_{1,2})e^{\lambda_1 t}\right)$$

$$\langle \Delta n_2(0)\Delta n_2(t)\rangle = \frac{\langle n_2\rangle^*}{\lambda_1-\lambda_2}\left((\lambda_1+q_{2,1}+q_{2,3})e^{\lambda_2 t}-(\lambda_2+q_{2,1}+q_{2,3})e^{\lambda_1 t}\right)$$

$$\langle \Delta n_1(0)\Delta n_2(t)\rangle = \frac{q_{1,2}\langle n_1\rangle^*}{\lambda_1-\lambda_2}\left(e^{\lambda_1 t}-e^{\lambda_2 t}\right)$$

$$\langle \Delta n_2(0)\Delta n_1(t)\rangle = \frac{q_{2,1}\langle n_2\rangle^*}{\lambda_1-\lambda_2}\left(e^{\lambda_1 t}-e^{\lambda_2 t}\right).$$

From the cross-correlation functions we get

$$\lim_{t\to 0}\frac{\langle \Delta n_1(0)\Delta n_2(t)\rangle - \langle \Delta n_2(0)\Delta n_1(t)\rangle}{t} = q_{1,2}\langle n_1\rangle^* - q_{2,1}\langle n_2\rangle^*.$$

The Equilibrium Grand Canonical Ensemble

Suppose the source and the sink are one and the same. Let state 0 be that state and suppose there are N internal states. Then the total number of particles in the system are given with probability

$$P(n) \propto \frac{Q(\beta, n)}{n!} e^{\beta \mu n},$$

The quantum-mechanical canonical partition function, $Q(\beta, n)$, is

$$Q(\beta, 1) \triangleq \sum_{i=0}^{N} e^{-\beta E_i}$$

for $n = 1$ and

$$Q(\beta, n) = (Q(\beta, 1))^n$$

for general n.

The Equilibrium Grand Canonical Ensemble

The normalizing condition for the grand canonical ensemble yields the grand canonical partition function

$$\begin{aligned}\Xi(\beta,\mu) &= \sum_{n=0}^{\infty} \frac{Q(\beta,n)}{n!} e^{\beta\mu n} \\ &= e^{Q(\beta,1)e^{\beta\mu}},\end{aligned}$$

$\Xi(\beta,\mu)$ is related to thermodynamics by $PV = k_B T \ln \Xi(\beta,\mu)$, from which the general differential $d(PV) = SdT + PdV + \langle n \rangle^{eq} d\mu$ gives

$$\begin{aligned}\langle n \rangle^{eq} &= Q(\beta,1)e^{\beta\mu} \\ \mu &= k_B T \ln \langle n \rangle^{eq} - k_B T \ln Q(\beta,1).\end{aligned}$$

The Equilibrium Grand Canonical Ensemble

The probability that there are n particles in the system is

$$P(n) = \frac{\left(Q(\beta,1)e^{\beta\mu}\right)^n}{n!}e^{-Q(\beta,1)e^{\beta\mu}} = \frac{\left(\langle n\rangle^{eq}\right)^n}{n!}e^{-\langle n\rangle^{eq}}$$

The grand canonical model gives

$$P(n_0, n_1, \ldots, n_N) = \prod_{i=0}^{N}\left[\frac{\left(\langle n_i\rangle^{eq}\right)^{n_i}}{n_i!}e^{-\langle n_i\rangle^{eq}}\right].$$

These results are related to each other by

$$P(n) = \sum_{\substack{n_0,n_1,\ldots,n_N\geq 0\\ n_0+n_1+\cdots+n_N=n}} P(n_0, n_1, \ldots, n_N) = \frac{\left(\langle n\rangle^{eq}\right)^n}{n!}e^{-\langle n\rangle^{eq}},$$

where $\langle n\rangle^{eq} = \langle n_0\rangle^{eq} + \langle n_1\rangle^{eq} + \cdots + \langle n_N\rangle^{eq}$.

The Kolmogorov Equations

From the Kolmogorov forward equation

$$\frac{d\langle n_i(t)\rangle}{dt} = \sum_{\substack{j=0 \\ j\neq i}}^{N+1} \left(\langle n_j(t)\rangle\, q_{j,i} - \langle n_i(t)\rangle\, q_{i,j}\right),$$

we get the associated backward equation

$$\frac{du_i(t)}{dt} = \sum_{\substack{j=0 \\ j\neq i}}^{N+1} \left[(u_j(t) - u_i(t))\, q_{i,j}\right] \qquad \forall i \in \{1, 2, \ldots, N\}$$

$$\text{ICs:} \quad u_i(0) = 0, \quad \forall i \in \{1, 2, \ldots, N\},$$

where $u_0(t) = 1$, $u_{N+1}(t) = \frac{n_{N+1}}{n_0}\frac{\pi_0}{\pi_{N+1}}$, and

$$u_i(t) = \frac{\langle n_i(t)\rangle}{n_0}\frac{\pi_0}{\pi_i}.$$

The Grand Canonical Model and Thermodynamics

The substitution allows for a physically meaningful potential function to be defined as

$$\mu_i(t) \triangleq \ln u_i(t) = -\ln \frac{\pi_i}{\pi_0} + \ln \langle n_i(t) \rangle - \ln n_0$$

and the chemical potential difference for a reaction between states i and j is

$$\Delta\mu_{i,j}(t) = \mu_i(t) - \mu_j(t) = \ln \frac{\langle n_i(t) \rangle \, q_{i,j}}{\langle n_j(t) \rangle \, q_{j,i}} = \ln \frac{J_{i,j}(t)}{J_{j,i}(t)}.$$

It follows that

$$\left(J_{i,j}(t) - J_{j,i}(t)\right) \Delta\mu_{i,j}(t) = \left(J_{i,j}(t) - J_{j,i}(t)\right) \ln \frac{J_{i,j}(t)}{J_{j,i}(t)} \geq 0,$$

which is equivalent to the second law of thermodynamics.

The Reaction Conductance

In terms of the chemical affinity, $u(t) = e^{\mu(t)}$, we have

$$\Delta u_{i,j}(t) = u_i(t) - u_j(t) = \frac{\langle n_i(t) \rangle \pi_0}{n_0 \pi_i} - \frac{\langle n_j(t) \rangle \pi_0}{n_0 \pi_j} = \frac{J_{i,j}(t) - J_{j,i}(t)}{n_0 \pi_i q_{i,j}} \pi_0$$

in which case

$$\left(J_{i,j}(t) - J_{j,i}(t)\right) \Delta u_{i,j}(t) = \frac{\left(J_{i,j}(t) - J_{j,i}(t)\right)^2}{n_0 \pi_i q_{i,j}} \pi_0 \geq 0.$$

This result bears a likeness to the linear Ohm's Law of electrical circuit theory. Considering this, a reaction conductance can be defined as

$$c_{i,j} \triangleq \frac{J_{i,j}(t) - J_{j,i}(t)}{\Delta u_{i,j}(t)} = \frac{n_0 \pi_i q_{i,j}}{\pi_0} = c_{j,i}.$$

Analogies to Random Walks

The system

$$\frac{du_i(t)}{dt} = \sum_{\substack{j=0 \\ j \neq i}}^{N+1} \left[\left(u_j(t) - u_i(t) \right) q_{i,j} \right] \qquad \forall i \in \{1, 2, \ldots, N\}$$

$$\text{ICs:} \quad u_i(0) = 0, \quad \forall i \in \{1, 2, \ldots, N\}$$

with $u_0(t) = 1$ and $u_{N+1}(t) = 0$ is equivalent to a random walk, where $u_i(t)$ is the expected number of times a walk, starting at state i at time $t = 0$, reaches state 0 before reaching state $N + 1$ and does so before time t has passed (Kelly, 1979).

Furthermore, when this system is in NESS, it is equivalent to an absorbing Markov chain and we can suggest a novel experimental method for measuring chemical affinities.

Discussion of the Grand Canonical Model

- The grand canonical model provides many important results that can be useful in studying reaction systems that actively exchange materials with their surroundings.
- The chemical species of a physical system may only be present in small numbers, which would negate the use of mass action kinetics.
- One of the most important results from this model is that the joint probability function of the number of molecules in each state can be found analytically.
- We are able to derive clear definitions of reaction potentials and conductances and suggest novel experimental methods to measure NESS fluxes and affinities.

Outline

- Background and motivation for the new modeling methods.
- The grand canonical model.
- Stochastic simulation algorithms.
- Stoichiometric constraints-based optimization approaches.
- Conclusions.

Stochastic Simulations

- Most biochemical networks are very complex and it is not possible to obtain analytic solutions when modeling them. For this reason, we turn to stochastic simulation algorithms.

Stochastic Simulations

- Most biochemical networks are very complex and it is not possible to obtain analytic solutions when modeling them. For this reason, we turn to stochastic simulation algorithms.

- Again, the mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.

Stochastic Simulations

- Most biochemical networks are very complex and it is not possible to obtain analytic solutions when modeling them. For this reason, we turn to stochastic simulation algorithms.
- Again, the mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.
- Gillespie (1976) developed an influential, exact method for simulating these networks.

Stochastic Simulations

- Most biochemical networks are very complex and it is not possible to obtain analytic solutions when modeling them. For this reason, we turn to stochastic simulation algorithms.
- Again, the mathematical basis for this approach is the discrete-state, continuous-time Markov process, also known as the (chemical) master equation approach.
- Gillespie (1976) developed an influential, exact method for simulating these networks.
- Other approximation methods have been developed to streamline the simulations and reduce computational overhead.

The Chemical Master Equation

The master equation is the time-evolution equation for the function $P(\mathbf{n}, t)$, where n_i is the number of molecules of species X_i in a well-mixed volume.

The Chemical Master Equation

The master equation is the time-evolution equation for the function $P(\mathbf{n}, t)$, where n_i is the number of molecules of species X_i in a well-mixed volume. If there are M different reactions (events) then we have

$$P(\mathbf{n}, t+dt) = P(\mathbf{n}, t)P(\text{there is no change within } dt)$$
$$+ \sum_{\mu=1}^{M} P(\mathbf{n} - \mathbf{s}_\mu, t)P(\text{reaction } \mu \text{ occurs within } dt),$$
$$= P(\mathbf{n}, t)\left(1 - \sum_{\mu=1}^{M} a_\mu(\mathbf{n})dt\right) + \sum_{\mu=1}^{M} P(\mathbf{n} - \mathbf{s}_\mu, t)a_\mu(\mathbf{n} - \mathbf{s}_\mu)dt,$$

where $a_\mu(\mathbf{n})dt$ is the probability that reaction μ will occur in $(t, t+dt)$ given that the system is in state $\mathbf{n}$ at time t and $\mathbf{s}_\mu$ is a stoichiometric vector defining the result of reaction μ.

The Chemical Master Equation

The master equation is the time-evolution equation for the function $P(\mathbf{n}, t)$, where n_i is the number of molecules of species X_i in a well-mixed volume. If there are M different reactions (events) then we have

$$P(\mathbf{n}, t+dt) = P(\mathbf{n}, t)P(\text{there is no change within } dt)$$
$$+ \sum_{\mu=1}^{M} P(\mathbf{n}-\mathbf{s}_\mu, t)P(\text{reaction } \mu \text{ occurs within } dt),$$
$$= P(\mathbf{n}, t)\left(1 - \sum_{\mu=1}^{M} a_\mu(\mathbf{n})dt\right) + \sum_{\mu=1}^{M} P(\mathbf{n}-\mathbf{s}_\mu, t)a_\mu(\mathbf{n}-\mathbf{s}_\mu)dt,$$

where $a_\mu(\mathbf{n})dt$ is the probability that reaction μ will occur in $(t, t+dt)$ given that the system is in state $\mathbf{n}$ at time t and $\mathbf{s}_\mu$ is a stoichiometric vector defining the result of reaction μ. From this we get

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_{\mu=1}^{M} a_\mu(\mathbf{n}-\mathbf{s}_\mu)P(\mathbf{n}-\mathbf{s}_\mu, t) - a_\mu(\mathbf{n})P(\mathbf{n}, t).$$

The Reaction Probability Density Function

So we ask: “When will the next event occur and what type of event will it be?”

The Reaction Probability Density Function

So we ask: “When will the next event occur and what type of event will it be?”

Define

$$P(\tau,\mu)d\tau \equiv \text{probability at time } t \text{ that the next event will occur in the differential time interval } (t+\tau, t+\tau+d\tau) \text{ and will be a type } \mu \text{ event,}$$

where $0 \leq \tau < \infty$ and μ simply indicates what type of event occurs.

The Reaction Probability Density Function

So we ask: "When will the next event occur and what type of event will it be?"

Define

$$P(\tau,\mu)d\tau \equiv \text{probability at time } t \text{ that the next event will occur in the differential time interval } (t+\tau, t+\tau+d\tau) \text{ and will be a type } \mu \text{ event,}$$

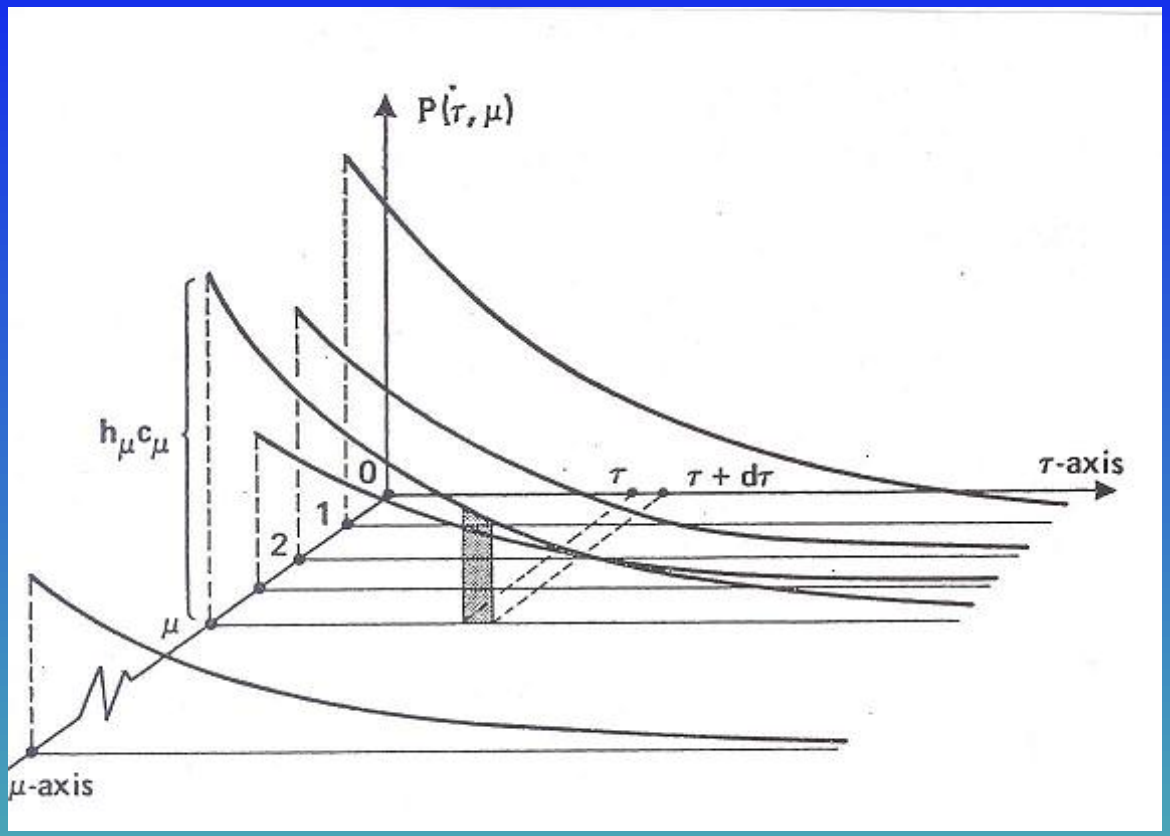
where $0 \leq \tau < \infty$ and μ simply indicates what type of event occurs.

This joint probability density function can be written as

$$P(\tau,\mu)d\tau = P_0(\tau)a_\mu d\tau = a_0 e^{-a_0\tau}\left(\frac{a_\mu}{a_0}\right)d\tau = P(\tau)P(\mu)d\tau$$

where $P_0(\tau)$ is the probability that no event occurs in the time interval $(t, t+\tau)$.

Schematic of the Density Function



(Gillespie, 1976)

The Gillespie Algorithm

$$P(\tau) = a_0 e^{-a_0 \tau} \quad \longrightarrow \quad F(\tau) = 1 - e^{-a_0 \tau}$$

$$P(\mu) = \frac{a_\mu}{a_0} \quad \longrightarrow \quad F(\mu) = \sum_{k=1}^{\mu} P(k)$$

The Gillespie Algorithm

$$P(\tau) = a_0 e^{-a_0 \tau} \quad \longrightarrow \quad F(\tau) = 1 - e^{-a_0 \tau}$$

$$P(\mu) = \frac{a_\mu}{a_0} \quad \longrightarrow \quad F(\mu) = \sum_{k=1}^{\mu} P(k)$$

So choose r_1 and r_2 from uniform distribution in the unit interval and

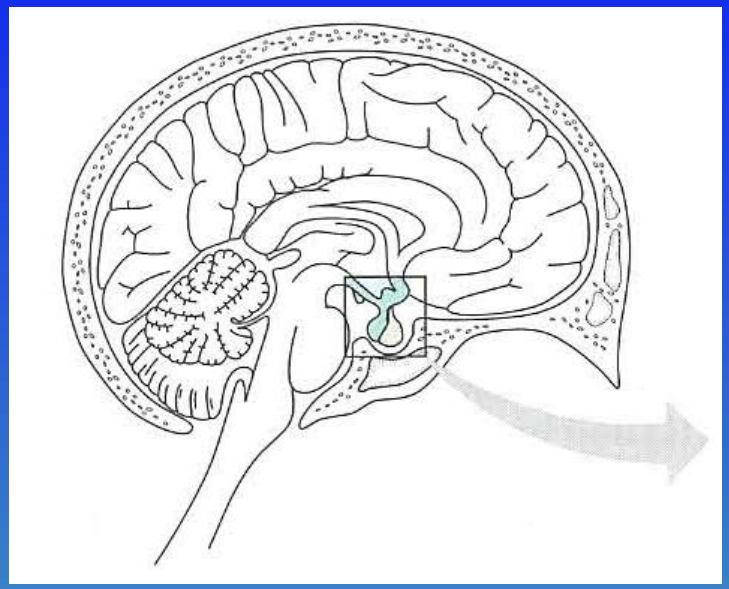
$$\tau = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right)$$

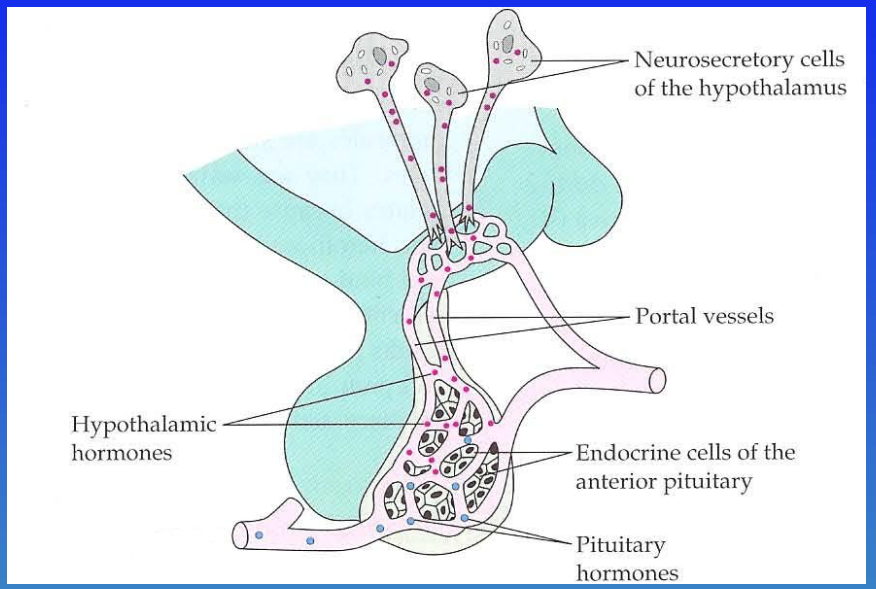
$$\sum_{k=1}^{\mu-1} \frac{a_k}{a_0} < r_2 \leq \sum_{k=1}^{\mu} \frac{a_k}{a_0}.$$

Testosterone in Men's Blood

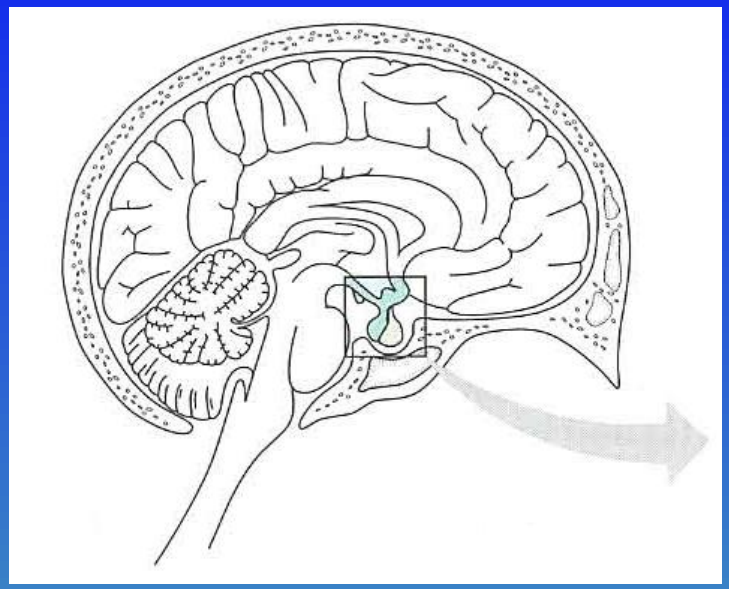
- Approximately 90% to 95% of testosterone in men is produced by the testes with typical blood testosterone levels in the range of 3 to 10 ng/mL.
- These levels have been experimentally observed to oscillate with a period of about 2 to 3 hours.
- An imbalance can cause dramatic changes (mood, acne, and weight).
- Pathway is associated with many other important processes in the body.
- Pharmaceutical interests in chemical castration (Goserelin, Lupron, and Depo-provera) and to create a male *pill*.

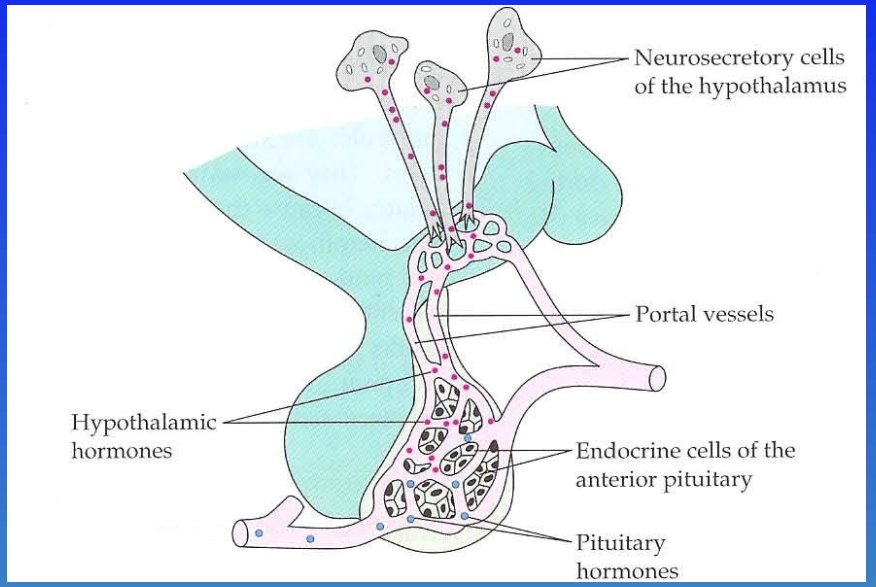
The Hypothalmus-Pituitary-Testicular Axis

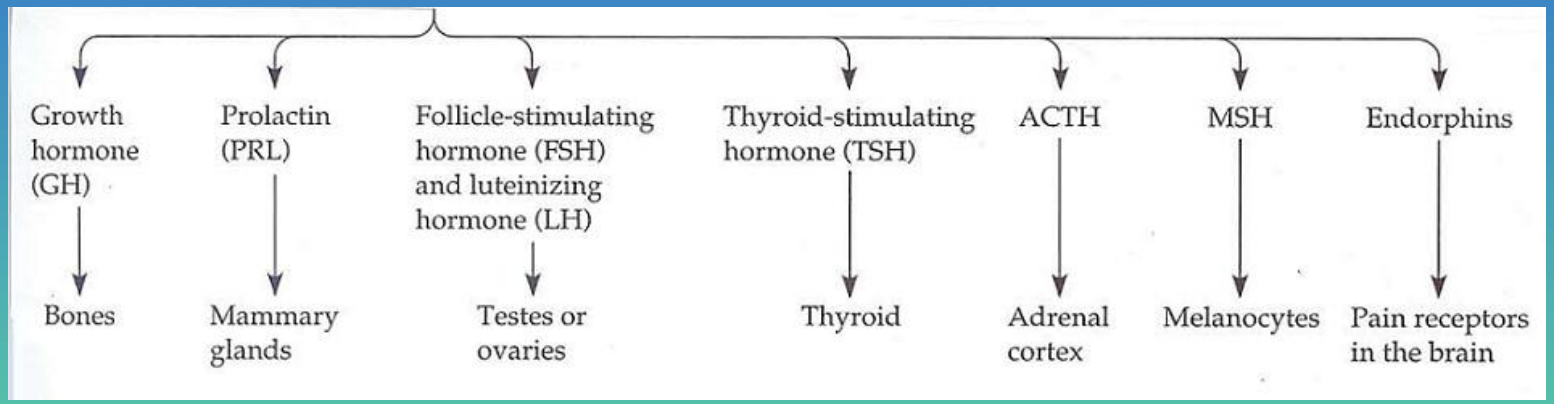




The Hypothalmus-Pituitary-Testicular Axis

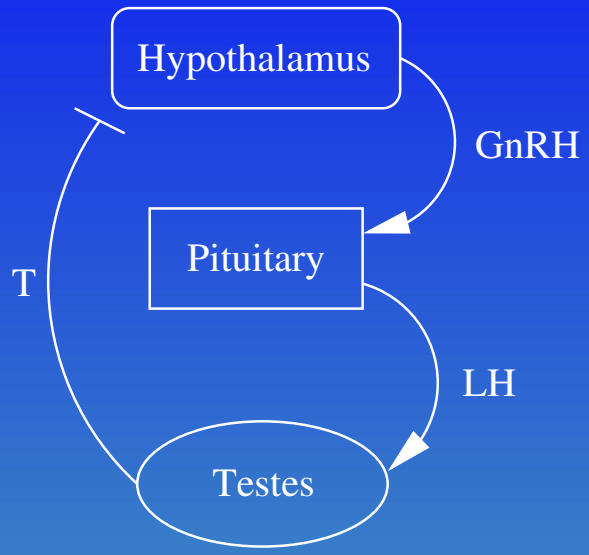






(Modified from Campbell, 1996.)

The Hormone Secretion Signaling System

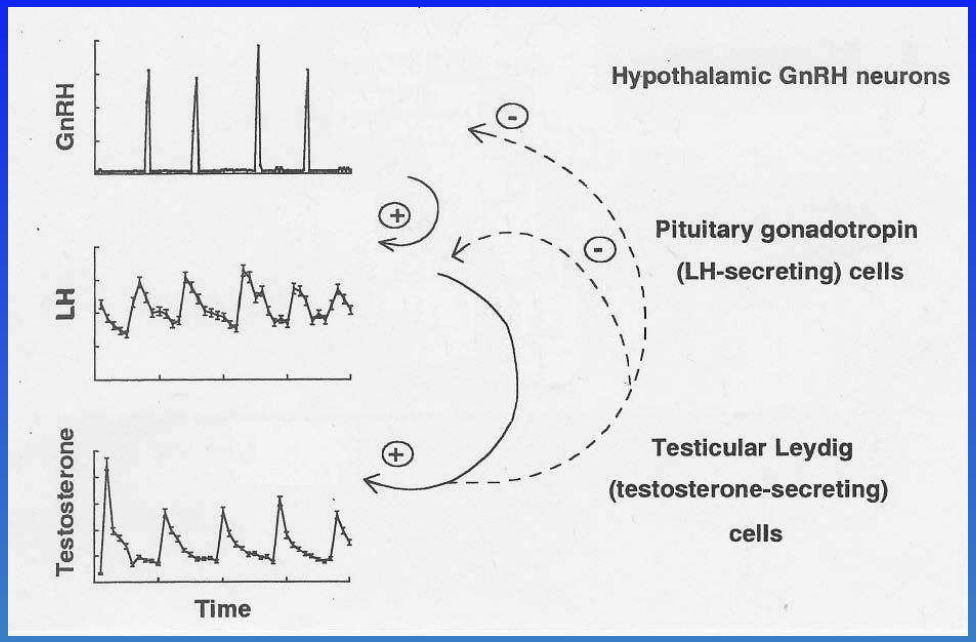


$GnRH =$ Gonadotropin Releasing Hormone

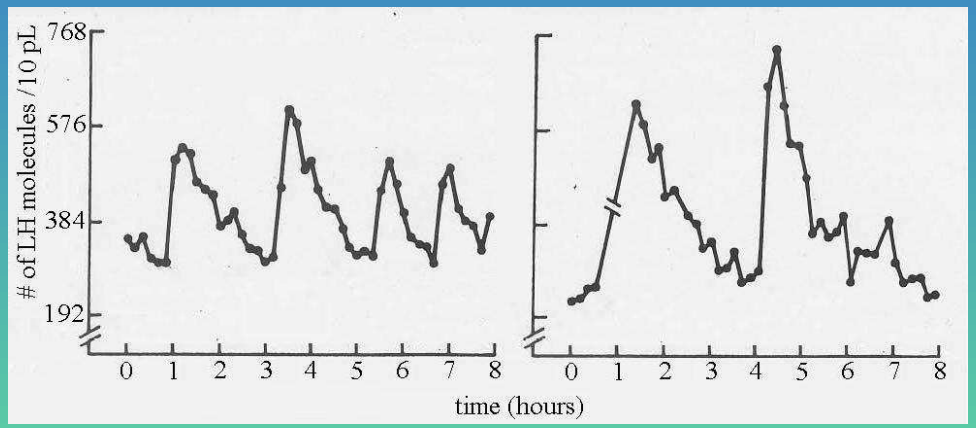
$LH =$ Luteinizing Hormone

$T =$ Testosterone

Experimental Observations



(Modified from Yen *et al.*, 1999.)



(Modified from Naftolin *et al.*, 1973.)

The Deterministic Model

If we represent the concentrations of *GnRH, LH*, and *T* by *R(t), L(t)*, and *T(t)*, respectively, then a proposed deterministic model of this system is

$$\frac{dR}{dt} = f(T) - b_1 R$$
$$\frac{dL}{dt} = g_1 R - b_2 L$$
$$\frac{dT}{dt} = g_2 L - b_3 T$$

The Deterministic Model

If we represent the concentrations of *GnRH, LH*, and *T* by *R(t), L(t)*, and *T(t)*, respectively, then a proposed deterministic model of this system is

$$\frac{dR}{dt} = f(T) - b_1 R$$
$$\frac{dL}{dt} = g_1 R - b_2 L$$
$$\frac{dT}{dt} = g_2 L - b_3 T$$

where

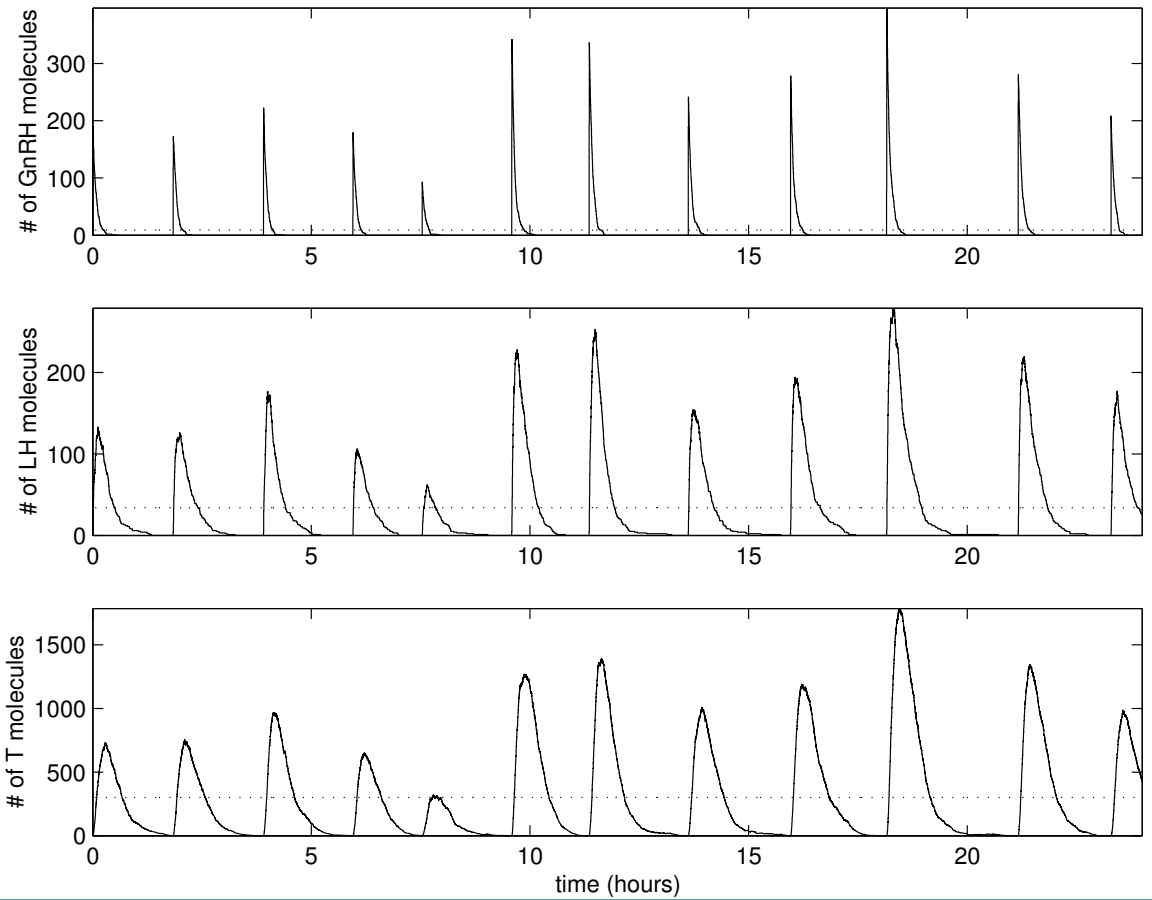
$$f(T) = \frac{A}{K + T}$$

and A, K, b_1, b_2, b_3, g_1, and g_2 are all positive constants.

History of the Model

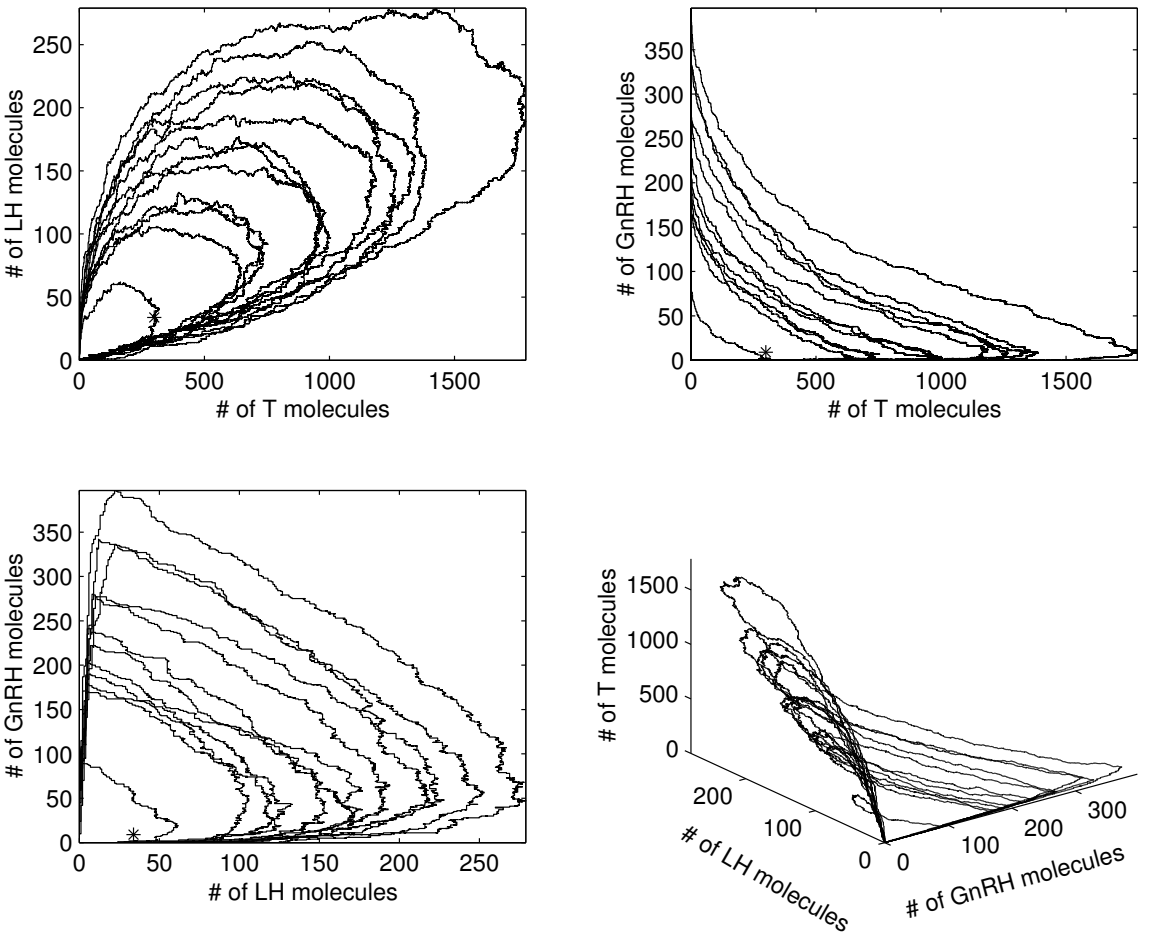
- Goodwin (1964) first proposed the model to demonstrate oscillatory behavior in enzymatic control processes.
- Smith (1980) studied slight variation involving a Hill coefficient in $f(T)$.
- Murray (1989) suggested using a time-delay in the production rate of T.
- Enciso and Sontag (2004) proved that the system has a globally stable fixed point (regardless of the length of the time-delay) and therefore does not have a limit cycle or sustained oscillations.
- More detailed (and more complicated) models include those by Cartwright and Husain (1986) and Keenan *et al.* (1998 and 2000).

A Stochastic Simulation of Hormone Secretion



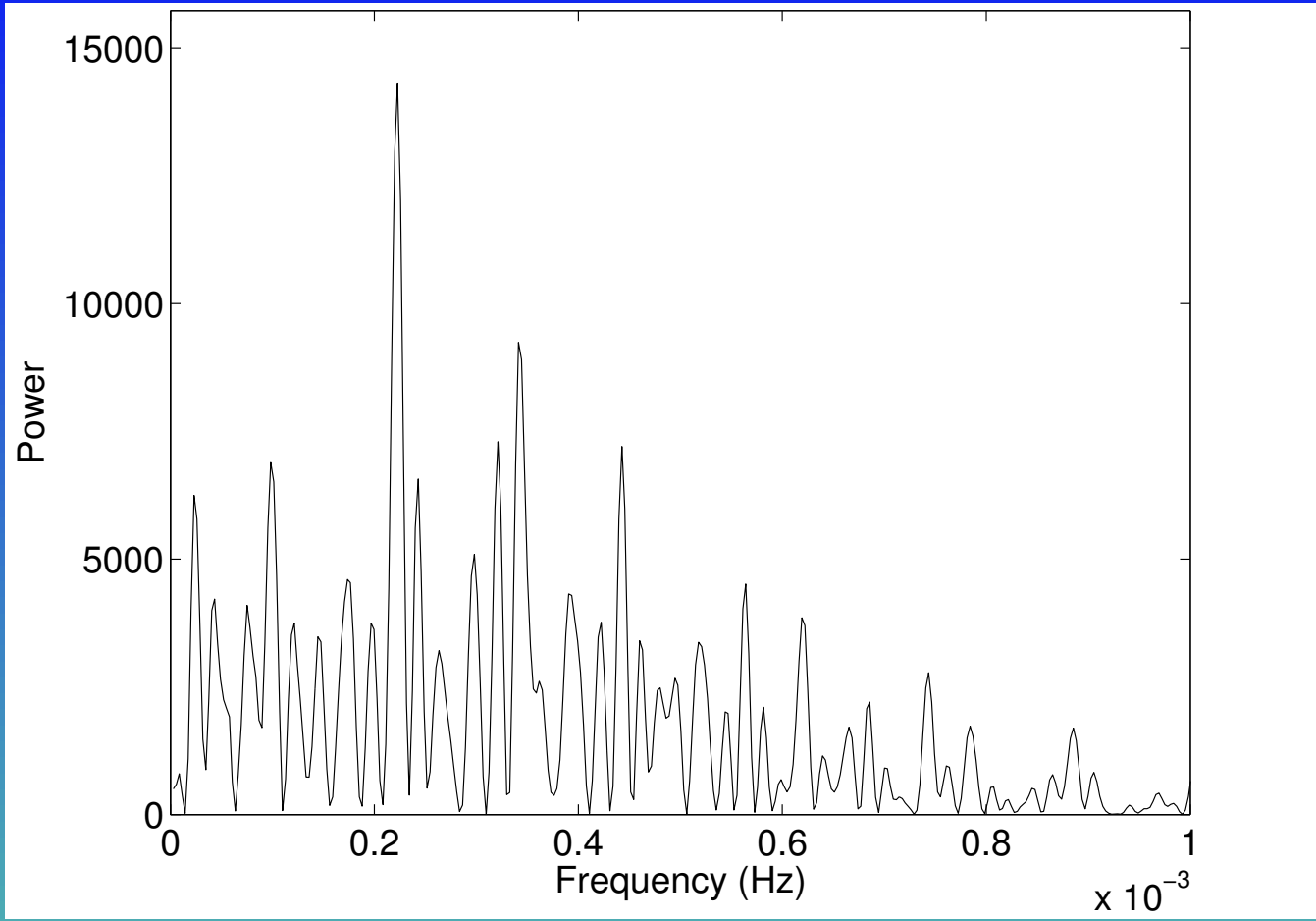
Simulation of hormone secretion for parameter values, $A = 10^{-4}$, $K = 10^{-7}$, $b_1 = 0.23$, $b_2 = 0.07$, $b_3 = 0.1$, $g_1 = 0.2618$, and $g_2 = 0.9015$. Average number of molecules are represented by dashed lines; average R is 9.09, average L is 33.92, and average T is 300.07.

A Stochastic Simulation of Hormone Secretion



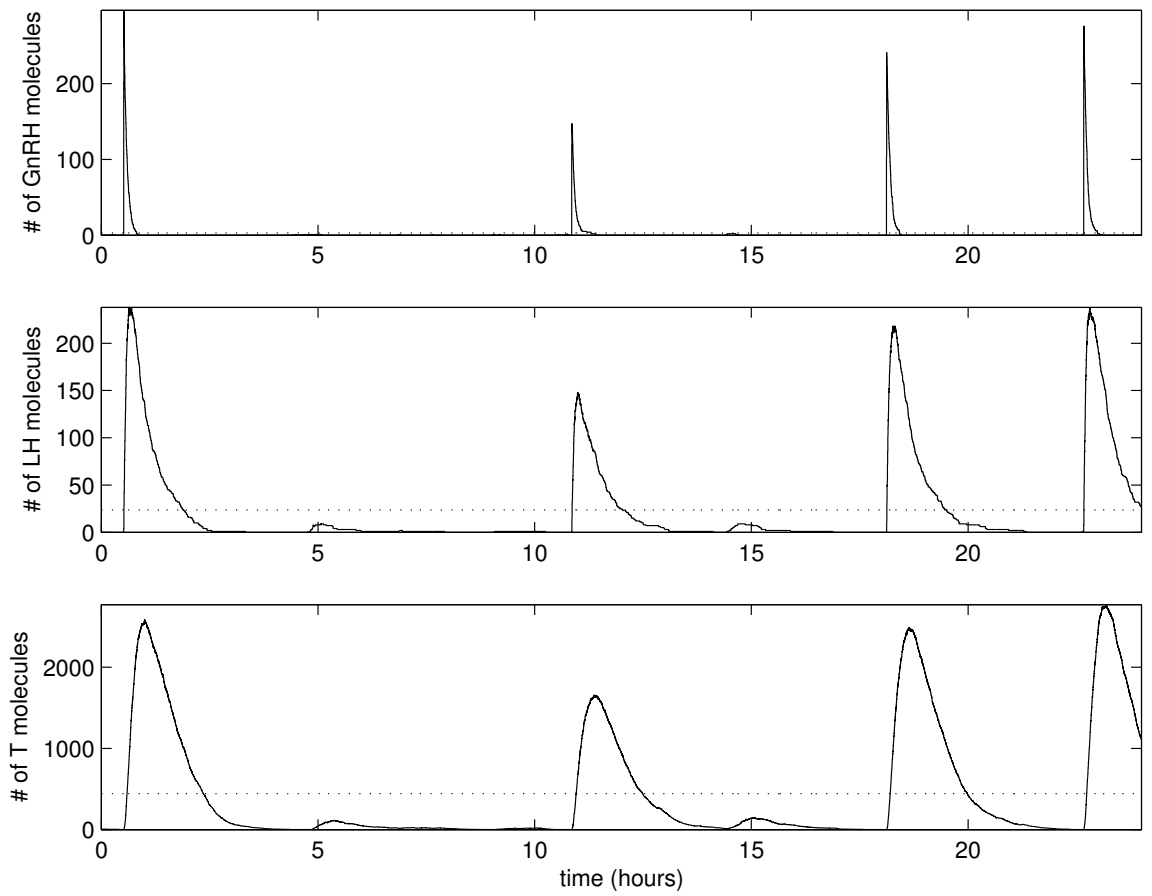
Two dimensional projections and three dimensional plot of simulation trajectory for the physical parameter values, $A = 10^{-4}$, $K = 10^{-7}$, $b_1 = 0.23$, $b_2 = 0.07$, $b_3 = 0.1$, $g_1 = 0.2618$, and $g_2 = 0.9015$. Average number of molecules are represented by asterisks; average R is 9.09, average L is 33.92, and average T is 300.07.

Lomb Spectral Analysis



The largest peak corresponds to a frequency of 2.3429×10^{-4} Hz.

The Switching Behavior



Simulation to illustrate the the switching behavior. Parameter values are $A = 10^{-1}$, $K = 10^{-4}$, $b_1 = 0.23$, $b_2 = 0.032$, $b_3 = 0.046$, $g_1 = 0.2618$, and $g_2 = 0.9015$.

Approximation Methods

- Gibson and Bruck (2000) proposed an approximation for systems in which some reactions occur much more often than others by reducing the number of random variables simulated.
- Gillespie (2001) introduced the τ-leap methods that make larger time steps and allow more events to occur within those steps as long as changes in the event probabilities stay within some tolerance.
- Burrage and Tian (2003) attempted to simulate continuous-time, continuous-state, stochastic-approximation, models driven by Wiener noise by introducing the framework of Poisson–Runge–Kutta methods.
- Turner, Schnell, and Burrage (2004) included fluctuations caused by the structural organisation of the cytoplasm and the limited diffusion of molecules due to macromolecular crowding.
- Burrage, Tian, and Burrage (2004) used multi-scale methods to incorporate the quasi-steady-state assumption with slow, intermediate, and fast reactions.

Discussion of Stochastic Simulations

- When we are interested in the effects of intrinsic fluctuations and are not able to obtain analytic results, we can rely on simulation methods such as the Gillespie algorithm.
- These algorithms provide a more realistic representation of a system than the deterministic, mass-action equations.
- By approaching the hormone model from a different physical basis we saw how intrinsic fluctuations can incite oscillations for low numbers of molecules.
- Even though the deterministic model has a globally stable fixed point, the stochastic model was able to capture the pulsatile behavior of the blood hormone levels.

Outline

- Background and motivation for the new modeling methods.
- The grand canonical model.
- Stochastic simulation algorithms.
- Stoichiometric constraints-based optimization approaches.
- Conclusions.

Stoichiometric Constraints-Based Approaches

- The mass-action models require detailed kinetic rate information, which, for the most part, is not available from experiment.

Stoichiometric Constraints-Based Approaches

- The mass-action models require detailed kinetic rate information, which, for the most part, is not available from experiment.
- To address this problem, genome-scale constraints-based models have been developed to describe the functional states, or phenotypes, of many organisms (Westerhoff and Palsson, 2004).

Stoichiometric Constraints-Based Approaches

- The mass-action models require detailed kinetic rate information, which, for the most part, is not available from experiment.
- To address this problem, genome-scale constraints-based models have been developed to describe the functional states, or phenotypes, of many organisms (Westerhoff and Palsson, 2004).
- Stoichiometric Network Theory (SNT) uses the static, algebraic structure of biochemical networks, within which chemical “motion” must take place.

Stoichiometric Constraints-Based Approaches

- The mass-action models require detailed kinetic rate information, which, for the most part, is not available from experiment.
- To address this problem, genome-scale constraints-based models have been developed to describe the functional states, or phenotypes, of many organisms (Westerhoff and Palsson, 2004).
- Stoichiometric Network Theory (SNT) uses the static, algebraic structure of biochemical networks, within which chemical “motion” must take place.
- This method of anlysis has been successfully applied to systems such as *E. coli* (Edwards and Palsson, 2000), mitochondrial energy metaboilsm (Ramakrishna et al., 2001), and metabolism in hepatocyte cells (Beard and Qian, 2005).

Stoichiometric Network Theory

Returning to the general mass-action equation for a system of N species and M reactions

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{M} (\kappa_i^j - \nu_i^j)(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}) + J_i^{ext}$$

we can rewrite the system of equations in matrix form as

$$\frac{d\mathbf{x}}{dt} = \mathbf{SJ} + \mathbf{J}^{ext}.$$

Since this system is being driven by external fluxes, it will go to a NESS.

Flux Balance Analysis

In NESS, the concentrations of the chemical species are not changing and we have

$$\mathbf{S}\mathbf{J} = -\mathbf{J}^{ext},$$

which is known as the flux balance constraint of FBA. Note that this constraint is similar to Kirchoff's current law of electrical circuit theory.

Additional constraints can be applied to the NESS fluxes such as

$$J^j_{lb} \leq J^j \leq J^j_{ub} \quad \forall j \in \{1, 2, \ldots, M\}$$
$$(J^{ext})^i_{lb} \leq (J^{ext})^i \leq (J^{ext})^i_{ub} \quad \forall i \in \{1, 2, \ldots, N\}.$$

Energy Balance Analysis

Define $\boldsymbol{\mu}$ as the N-dimensional vector of chemical potentials, then the M-dimensional vector of reaction potentials, $\Delta\boldsymbol{\mu}$, is given by

$$\mathsf{S}^T\boldsymbol{\mu} = \Delta\boldsymbol{\mu}.$$

We can define the nullspace matrix $\mathsf{K} \in \mathbb{R}^{M\times(M-N-r)}$ with columns that form a basis for the nullspace of S, so that $\mathsf{SK} = 0$. Then we have the constraint

$$\mathsf{K}^T\mathsf{S}^T\boldsymbol{\mu} = \mathsf{K}^T\Delta\boldsymbol{\mu} = \mathbf{0},$$

which is a constraint for the conservation of energy and is similar to Kirchoff's loop or voltage law of electrical circuit theory.

Energy Balance Analysis

If we define the nonnegative forward and reverse reaction fluxes so that $\mathbf{J} = \mathbf{J}_+ - \mathbf{J}_-$, then the reaction potential is

$$\Delta\mu^j = RT \ln\left(\frac{J_-^j}{J_+^j}\right),$$

which leads us directly to the second law of thermodynamics, i.e.,

$$-J^j \Delta\mu^j = -RT\left(J_+^j - J_-^j\right) \ln\left(\frac{J_-^j}{J_+^j}\right) \geq 0.$$

Entropy must increase and the system must dissipate heat,

$$hdr = -\mathbf{J}^T \Delta\boldsymbol{\mu} > 0.$$

The Optimization Problem

$$\min_{\mathbf{J},\mathbf{J}_+,\mathbf{J}_-,\mathbf{J}^{ext},\Delta\boldsymbol{\mu}} \quad f(\mathbf{J},\mathbf{J}_+,\mathbf{J}_-,\mathbf{J}^{ext},\Delta\boldsymbol{\mu})$$

s.t.

$$\mathsf{S}\mathbf{J} + \mathbf{J}^{ext} = \mathbf{0}$$

$$\mathsf{K}^T \Delta\boldsymbol{\mu} = \mathbf{0}$$

$$\text{diag}\left(e^{\Delta\boldsymbol{\mu}/RT}\right)\mathbf{J}_+ - \mathbf{J}_- = \mathbf{0}$$

$$\mathbf{J} - \mathbf{J}_+ + \mathbf{J}_- = \mathbf{0}$$

$$\mathbf{J}_{lb} \leq \mathbf{J} \leq \mathbf{J}_{ub}$$

$$\mathbf{0} \leq \mathbf{J}_+ < \boldsymbol{\infty}$$

$$\mathbf{0} \leq \mathbf{J}_- < \boldsymbol{\infty}$$

$$\mathbf{J}^{ext}_{lb} \leq \mathbf{J}^{ext} \leq \mathbf{J}^{ext}_{ub}$$

$$\Delta\boldsymbol{\mu}_{lb} \leq \Delta\boldsymbol{\mu} \leq \Delta\boldsymbol{\mu}_{ub}$$

Sequential Quadratic Programming

- We can solve the problem for any given, smooth, linear or nonlinear, objective function using a Sequential Quadratic Programming (SQP) algorithm.
- The basic idea of an SQP method is to step toward an optimal solution by iteratively approximating the problem by quadratic subproblems.
- A simple interpretation of an SQP algorithm is to view it as an application of Newton's method to the Karush–Kuhn–Tucker optimality conditions, i.e.,

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \nabla f(\mathbf{x}^*) - \sum_{i \in \mathcal{A}(\mathbf{x}^*)} \lambda_i^* \nabla c_i(\mathbf{x}^*) = \mathbf{0}.$$

The Quadratic Subproblem

Linearizing at the current iterate $\mathbf{x}_k$, we get the subproblem

$$\begin{aligned} \min_{\mathbf{p}} \quad & \frac{1}{2}\mathbf{p}^T \mathsf{H}_k \mathbf{p} + \nabla f_k^T \mathbf{p} \\ \text{s.t.} \quad & \nabla c_i(\mathbf{x}_k)^T \mathbf{p} + c_i(\mathbf{x}_k) = \mathbf{0}, \quad i \in \{1, 2, \ldots, m\} \\ & \nabla c_i(\mathbf{x}_k)^T \mathbf{p} + c_i(\mathbf{x}_k) \geq \mathbf{0}, \quad i \in \{m+1, \ldots, n\}, \end{aligned}$$

which gives the search direction used to update the current iterate

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \, \mathbf{p}_k$$

by doing a line search.

A Hypothetical Example

Consider

$$
\begin{aligned}
A + 2B &\rightleftharpoons C \\
C + D &\rightleftharpoons 2A + 2B \\
A + B &\rightleftharpoons 2D \\
A + C &\rightleftharpoons B + 3D \\
B &\rightleftharpoons D,
\end{aligned}
$$

for which

$$
\mathsf{S} = \begin{pmatrix} -1 & 2 & -1 & -1 & 0 \\ -2 & 2 & -1 & 1 & -1 \\ 1 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & 3 & 1 \end{pmatrix} \quad \text{and} \quad \mathsf{K} = \begin{pmatrix} -0.7163 & -0.3345 \\ -0.3205 & -0.4347 \\ 0.4710 & -0.6349 \\ -0.3958 & 0.1001 \\ -0.0752 & 0.5348 \end{pmatrix}
$$

A Hypothetical Example

Using FBA alone to maximize D output:

Case 1				Case 2			
rxn	$\mathbf{J}$	species	$\mathbf{J}^e$	rxn	$\mathbf{J}$	species	$\mathbf{J}^e$
1	0	A	1	1	0.05	A	10
2	0	B	0	2	10.09	B	20
3	1	C	0	3	20.17	C	20
4	0	D	-1	4	9.96	D	-90
5	-1			5	29.87		

A Hypothetical Example

Using FBA and EBA to maximize D output and minimize total energy:

rxn	$\mathbf{J}$	$\mathbf{J}_+$	$\mathbf{J}_-$	$\Delta\boldsymbol{\mu}$	$\mathbf{J}^e$	species
1	-0.95	13.63	14.58	0.067	10	A
2	8.35	56.04	47.69	-0.161	20	B
3	16.94	113.85	96.91	-0.161	20	C
4	10.70	85.22	74.52	-0.134	-90	D
5	32.36	143.66	111.30	-0.255	$hdr = 13.83$	

A Hypothetical Example

Using FBA, EBA, and heat constraint to maximize D and minimize total energy:

rxn	$\mathbf{J}$	$\mathbf{J}_+$	$\mathbf{J}_-$	$\Delta\boldsymbol{\mu}$	$\mathbf{J}^e$	species
1	0.00	6.04	6.04	0.00	10	A
2	10.44	12.87	2.43	-1.67	20	B
3	21.33	22.12	0.79	-3.33	20	C
4	9.56	11.78	2.23	-1.67	-90	D
5	29.11	29.31	0.20	-5.00	$hdr = 250$	

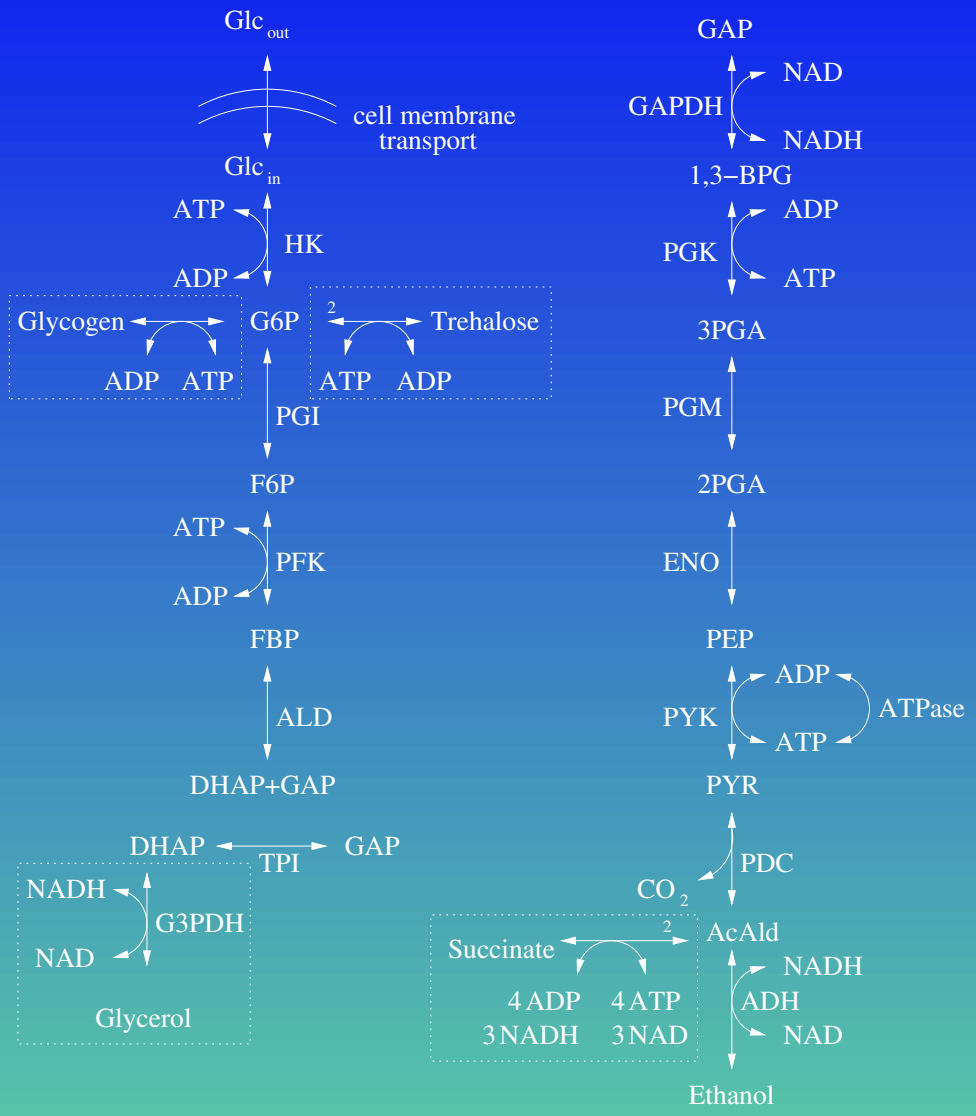
Saccharomyces cerevisiae

Overall, the net reaction of fermentation is the convertion of glucose to ethanol and carbon dioxide

$$C_6H_{12}0_6 \longrightarrow 2C_2H_5OH + 2CO_2.$$

- Under anaerobic conditions, most of the energy from the sugar is transferred to ethanol and growth of the yeast cells is minimized.
- Temperature is an important environmental factor for yeast because above the optimal temperature of 30°C, metabolism begins to slow, and when heat begins to denature the proteins in the cell, metabolism decreases rapidly.
- Under anaerobic conditions with a complex medium and glucose as the substrate, a continuous culture of *S. cerevisiae* has a specific rate of heat production of 0.2 $W \cdot g^{-1}$.

Saccharomyces cerevisiae



Saccharomyces cerevisiae

rxn	transport	HK	PGI	PFK	ALD	TPI	GAPDH
V_{min}							-24.3
V_{max}	0.36	0.84	1.26	0.68	1.19	8.4	4.4
rxn	PGK	PGM	ENO	PYK	PDC	ADH	
V_{min}	-4.8					-3.0	
V_{max}		9.4	1.35	4.05	0.65		

(Teusink et al., 2000)

Reaction fluxes listed have units of $(\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1})$ and potentials have units of $(\text{J} \cdot \text{mol}^{-1})$ and it is assumed that the temperature of the system is 30°C.

	unbranched					branched				
	FBA	FBA, EBA, & heat				FBA	FBA, EBA, & heat			
rxn	$\mathbf{J}$	$\mathbf{J}$	$\mathbf{J}_+$	$\mathbf{J}_-$	$\Delta\boldsymbol{\mu}$	$\mathbf{J}$	$\mathbf{J}$	$\mathbf{J}_+$	$\mathbf{J}_-$	$\Delta\boldsymbol{\mu}$
transport	0.33	0.33	1.02	0.69	-971	0.36	0.36	0.75	0.39	-1659
HK	0.33	0.33	1.02	0.69	-971	0.36	0.36	0.75	0.39	-1658
PGI	0.33	0.33	1.02	0.69	-971	0.30	0.30	0.70	0.40	-1420
PFK	0.33	0.33	1.02	0.69	-971	0.30	0.30	0.70	0.40	-1422
ALD	0.33	0.33	1.02	0.69	-971	0.30	0.30	0.70	0.40	-1420
TPI	0.33	0.33	1.02	0.69	-971	0.23	0.23	0.64	0.41	-1126
GAPDH	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2379
PGK	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2379
PGM	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2380
ENO	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2378
PYK	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2377
PDC	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2377
ADH	0.65	0.65	1.21	0.56	-1943	0.51	0.51	0.85	0.35	-2270
ATPase	0.65	0.65	1.21	0.56	-1943	0.32	0.31	0.72	0.41	-1422
Glycogen	-	-	-	-	-	0.02	0.02	0.25	0.23	-234
Trehalose	-	-	-	-	-	0.02	0.02	0.23	0.21	-203
Glycerol	-	-	-	-	-	0.07	0.07	0.41	0.34	-476
Succinate	-	-	-	-	-	0.01	0.01	0.22	0.20	-166

Discussion of Stoichiometric Constraints-Based Approaches

- SNT has been shown to be a very accurate and useful tool for studying mutant and disease affected organisms.
- By combining FBA and EBA constraints, we are certain that the feasible solutions are mass balanced and thermodynamically realistic.
- Using an SQP to solve the optimization problem allows us to combine the FBA and EBA constraints and consider objective functions many different objective functions.
- This method allows us to study a system on the whole genome scale and do *in silico* experiments instead of *in vitro* or *in vivo* experiments.

Outline

- Background and motivation for the new modeling methods.
- The grand canonical model.
- Stochastic simulation algorithms.
- Stoichiometric constraints-based optimization approaches.
- Conclusions.

Conclusions

- The classical methods for modeling biochemical networks are limited in their power.

Conclusions

- The classical methods for modeling biochemical networks are limited in their power.
- The grand canonical model is able to capture intrinsic fluctuations and provide analytic results for linear networks.

Conclusions

- The classical methods for modeling biochemical networks are limited in their power.
- The grand canonical model is able to capture intrinsic fluctuations and provide analytic results for linear networks.
- Stochastic simulation approaches are able to capture intrinsic fluctuations for more complicated systems, but no analytic results are available.

Conclusions

- The classical methods for modeling biochemical networks are limited in their power.
- The grand canonical model is able to capture intrinsic fluctuations and provide analytic results for linear networks.
- Stochastic simulation approaches are able to capture intrinsic fluctuations for more complicated systems, but no analytic results are available.
- Using stoichiometric constraints-based approaches, we are able to quantitatively study the possible phenotypes of a system.

Conclusions

- The classical methods for modeling biochemical networks are limited in their power.
- The grand canonical model is able to capture intrinsic fluctuations and provide analytic results for linear networks.
- Stochastic simulation approaches are able to capture intrinsic fluctuations for more complicated systems, but no analytic results are available.
- Using stoichiometric constraints-based approaches, we are able to quantitatively study the possible phenotypes of a system.
- It is clear that these new methods lead us toward to ultimate goal of developing a complete model of a living organism.

Thank You

- Hong Qian, my advisor.
- James Burke, Mark Kot, and Dan Beard.
- UW Department of Applied Math.
- NSF and DoD.